2013

Uncertainty Quantification using Exponential Epi-Splines, Proceedings of the International Conference on Structural Safety and Reliability, New York, NY.

Uncertainty Quantification using Epi-Splines and Soft Information

Johannes O. Royset
Visiting Professor, Mathematics, University of California, Davis
Associate Professor, OR, Naval Postgraduate School, Monterey

with N. Sukumar (CEE, UCD) and R. Wets (Math, UCD)

Stanford June 2013

This material is based upon work supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under grant numbers 00101-80683, W911NF-10-1-0246 and W911NF-12-1-0273.

Uncertainty quantification (UQ)

Engineering, biological, physical systems

- Input: random vector $\mathbf{V}$ ("known" distribution)
- System function G; implicitly defined e.g. by simulation
- Output: random variable

$$X = G(\mathbf{V})$$

Uncertainty quantification (UQ)

Engineering, biological, physical systems

- Input: random vector $\mathbf{V}$ ("known" distribution)
- System function G; implicitly defined e.g. by simulation
- Output: random variable

$$X = G(\mathbf{V})$$

Given observations (data) $x^1 = G(\mathbf{v}^1), ..., x^n = G(\mathbf{v}^n)$, we seek a description of X:

- mean, standard deviation
- quantile, superquantile
- distribution, density (pdf)

Main challenge: few data points; little relevant data
But soft information might be available

Approaches to UQ

Main approaches to characterize $X = G(\mathbf{V})$:

- Statistical estimation:
  - Sample $\mathbf{v}^1, \mathbf{v}^2, ..., \mathbf{v}^n$ from input distribution
  - Evaluate G to generate $x^i = G(\mathbf{v}^i)$, $i = 1, 2, ..., n$
  - Justification through limit laws of statistics as sample size becomes **large**

Approaches to UQ

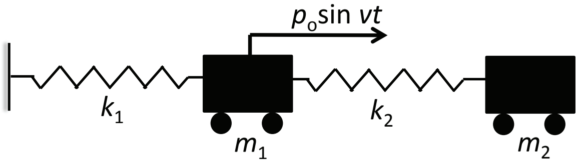
Main approaches to characterize $X = G(\mathbf{V})$:

- Statistical estimation:
  - Sample $\mathbf{v}^1, \mathbf{v}^2, ..., \mathbf{v}^n$ from input distribution
  - Evaluate G to generate $x^i = G(\mathbf{v}^i)$, $i = 1, 2, ..., n$
  - Justification through limit laws of statistics as sample size becomes **large**
- Expansion method:
  - Approximation of G using basis functions ϕ_j (polynomials):

$$\hat{G}(\mathbf{v}) = \sum_{j=1}^{k} c_j \phi_j(\mathbf{v})$$

  - By construction, easy to compute moments of $\hat{X} = \hat{G}(\mathbf{V})$
  - Exponential rate of convergence of $\hat{G}$ to G under smoothness; difficulties when **v high-dimensional**, G **nonsmooth**

Example: 2-dof dynamical system



$$m_1 \ddot{u}_1(t) + (k_1 + k_2) u_1(t) - k_2 u_2(t) = p_o \sin vt$$
$$m_2 \ddot{u}_2(t) - k_2 u_1(t) + k_2 u_2(t) = 0$$

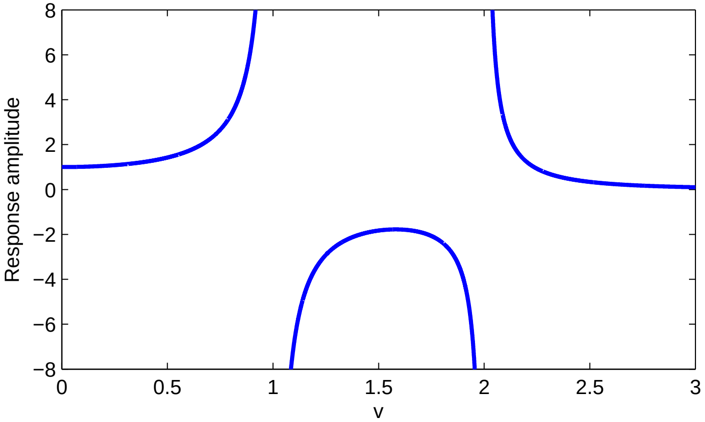
For choices of k_i, m_i, steady-state displacement at node 2:

$$u_2(t) = u_{2o} \sin vt \text{ with } u_{2o} = \frac{1}{(1 - v^2)(1 - v^2/4)}$$

Example: 2-dof dynamical system

For random input frequency V, get random response

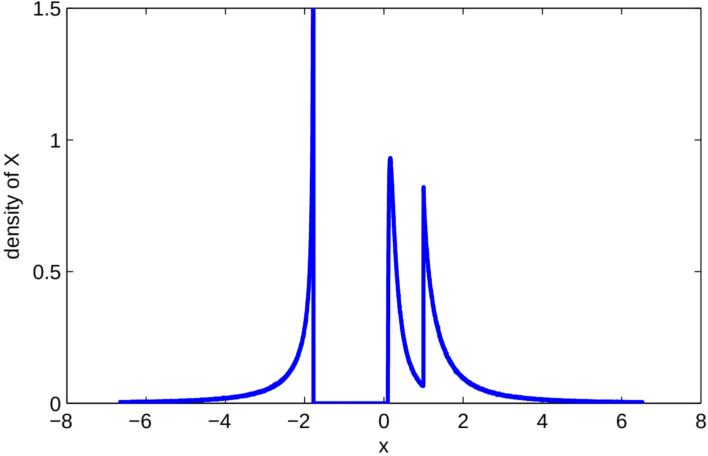
$$X = G(V) = 1/((1 - V^2)(1 - V^2/4))$$



G discontinuous; hopeless to approximate by polynomials

Example: 2-dof dynamical system

V mix of beta densities gives response density:



How to estimate densities like this with a small sample?

Our approach to UQ: density estimation

Random variable X: unknown distribution

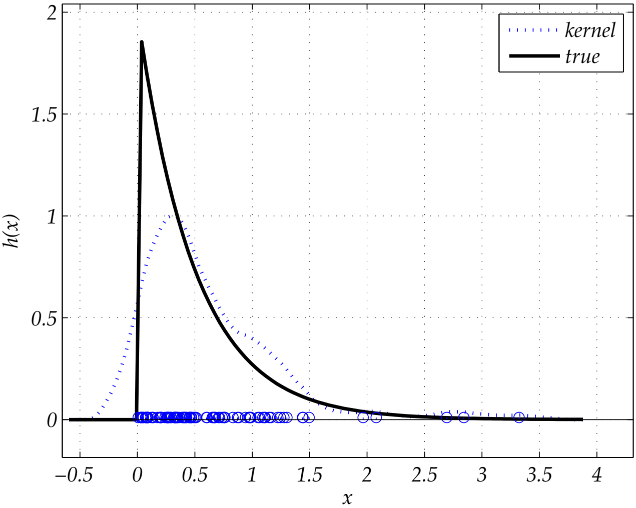
- Would like to estimate its probability density function h
- Have data in the form of random sample

$$X^1, X^2, ..., X^n$$

- Sample often independent and identically distributed as X, but sometimes not
- Parametric, Bayesian, Nonparametric approaches: Incorporate soft information to various degrees

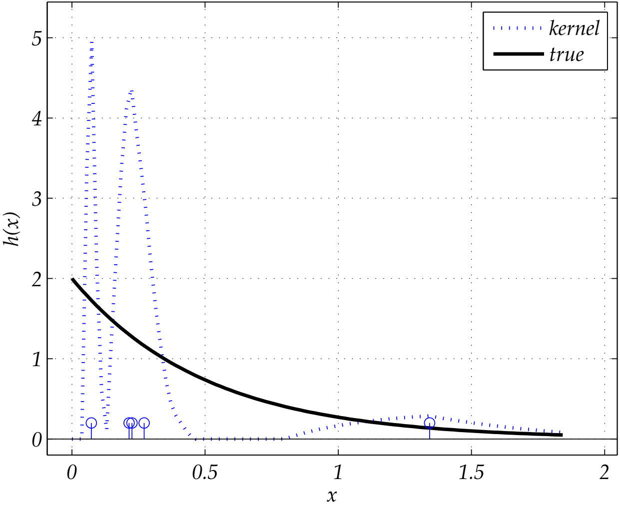
Example: Nonparametric density estimation

Kernel estimate with 100 sample points from exponential distribution



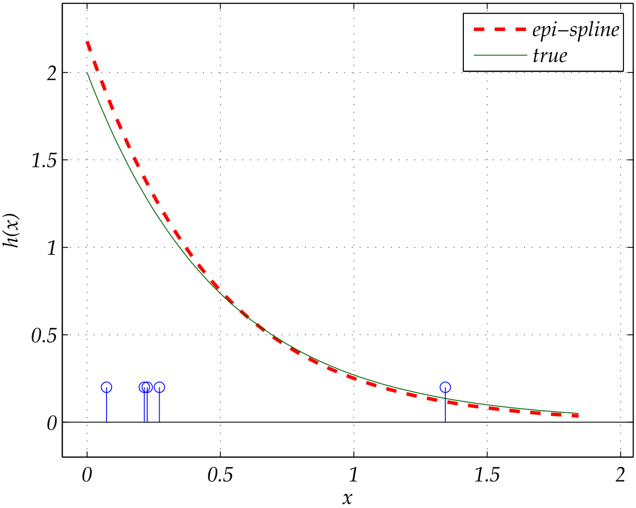
Example with small sample size

... and with 5 points



Using our epi-spline framework

Same 5 points, but with soft information (continuously differentiable, nonnegativity, and decreasing)



Our framework for UQ and density estimation

- Nonparametric density estimation utilizing soft information
- Includes essentially any soft information that we may have about X:
  - support bounds
  - density continuity, smoothness
  - density shape (unimodal, decreasing, etc.)
  - moments
  - proximity to known density
  - system knowledge (convex G, gradient of G)
- Span existing approaches: Bayesian, parametric, nonparametric

Using a maximum likelihood criterion, approach results in infinite-dimensional optimization problems

The optimization problem

- Maximizing log-likelihood function of sample X^1, ..., X^n
- subject to constraints derived from the soft information

$$
\begin{array}{rrcl}
h^n \in \underset{h}{\operatorname{argmax}} & \log\left(\prod_{i=1}^{n} h(X^i)\right)^{1/n} & = & \frac{1}{n}\sum_{i=1}^{n} \log h(X^i) \\
\text{s.t.} & \int_{-\infty}^{\infty} h(x)dx & = & 1 \\
& h & \geq & 0 \\
& h & \in & \mathcal{H}^n
\end{array}
$$

$\mathcal{H}^n \subset$ some function space; incorporates soft information

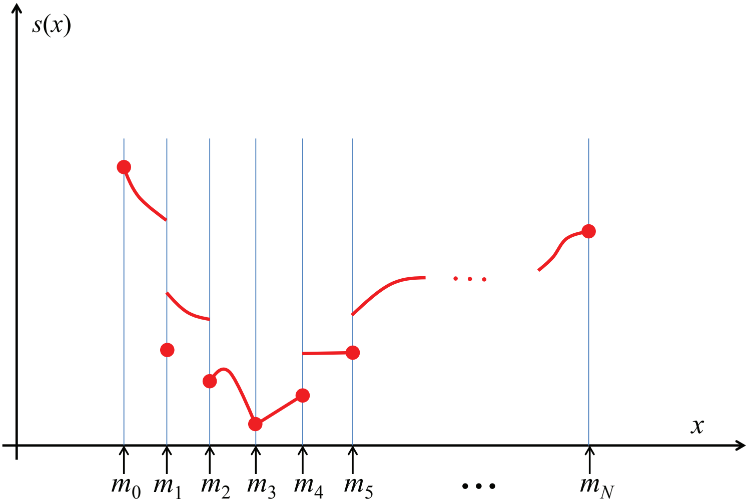
Exponential epi-spline estimator

Given sample $X^1, ..., X^n$, the exponential epi-spline estimator of the true density is

$$h^n(x) = e^{-s^n(x)},$$

where $s^n(x)$ is an epi-spline

Epi-splines: piecewise polynomial functions



Exponential epi-spline estimator

Given sample $X^1, ..., X^n$, the exponential epi-spline estimator of the true density is

$$h^n(x) = e^{-s^n(x)},$$

where $s^n(x)$ is an epi-spline

Main features:

- finite number of parameters; powerful optimization technology available
- approximates to arbitrary accuracy essentially any function
- easily includes soft information
- more flexible than 'classical' splines
- nonnegativity achieved automatically

Representation of epi-spline

Every $s \in \text{e-spl}^p(m)$, with $m = \{m_k\}_{k=0}^N$, is uniquely represented by

$$\mathbf{r} = (s_0, s_1, ..., s_N, \mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_N), \quad s_k \in R,\ \mathbf{a}_k \in R^{p+1},$$

such that

$$s(x) = \langle \mathbf{c}(x), \mathbf{r} \rangle,$$

where

$$\mathbf{c}(x) = \begin{cases} (0, ..., 0, 1, x - m_{k-1},\ ..., & (x - m_{k-1})^p, 0, ..., 0) \\ & \text{if } x \in (m_{k-1}, m_k), k = 1, ..., N \\ (0, ..., 0, 1, 0, ..., 0) & \text{if } x = m_k, k = 0, 1, ..., N. \end{cases}$$

Examples

- normal density represented by $\text{e-spl}^2(m)$ on $[m_0, m_N]$ for any m

Examples

- normal density represented by $\mathrm{e\text{-}spl}^2(m)$ on $[m_0, m_N]$ for any m
- exponential density represented by $\mathrm{e\text{-}spl}^1(m)$ on $[m_0, m_N]$ for any choice of m with $m_0 = 0$

Examples

- normal density represented by $\text{e-spl}^2(m)$ on $[m_0, m_N]$ for any m
- exponential density represented by $\text{e-spl}^1(m)$ on $[m_0, m_N]$ for any choice of m with $m_0 = 0$
- lognormal and Pareto also exactly represented after transformation

Examples

- normal density represented by $\text{e-spl}^2(m)$ on $[m_0, m_N]$ for any m
- exponential density represented by $\text{e-spl}^1(m)$ on $[m_0, m_N]$ for any choice of m with $m_0 = 0$
- lognormal and Pareto also exactly represented after transformation
- other densities approximated to arbitrary accuracy

Recall: maximize log-likelihood

$$
\begin{aligned}
h^n \in \operatorname*{argmax}_h \frac{1}{n}\sum_{i=1}^{n} \log h(X^i) \\
\text{s.t.} \qquad \int_{-\infty}^{\infty} h(x)dx &= 1 \\
h &\geq 0 \\
h &\in \mathcal{H}^n
\end{aligned}
$$

Now use $h(x) = e^{-\langle \mathbf{c}(x), \mathbf{r} \rangle}$ and optimize over vector $\mathbf{r}$ instead

Resulting optimization problem

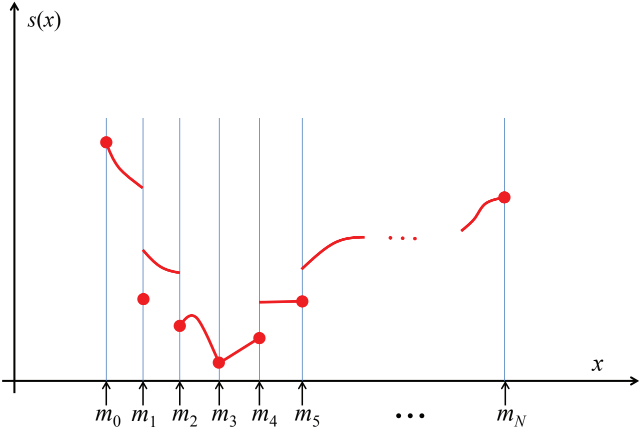
$$
\begin{array}{lrcl}
 & \min_{\mathbf{r}} \frac{1}{n} \sum_{i=1}^{n} \langle \mathbf{c}(X^i), \mathbf{r} \rangle & & \\
\text{s.t.} & \int_{m_0}^{m_N} e^{-\langle \mathbf{c}(x), \mathbf{r} \rangle} dx & = & 1 \\
 & r & \in & R
\end{array}
$$

R often polyhedral; $\leq$ often replaces $=$

$\implies$ convex problem

Formulation of soft information

Easy to ensure bounds on domain, continuity, smoothness, monotonicity



Formulation of soft information

Recall: Kullback-Leibler divergence from density h to density g

$$d_{KL}(h||g) = \int_{-\infty}^{\infty} h(x) \log \frac{h(x)}{g(x)} dx$$

If $s \in \text{e-spl}^p(m)$ and $\mathbf{r}$ its epi-spline parameter, then

$$d_{KL}(h||e^{-s(\cdot)}) = \left\langle \int_{m_0}^{m_N} \mathbf{c}(x)h(x)dx, \mathbf{r} \right\rangle + \int_{-\infty}^{\infty} (\log h(x))h(x)dx,$$

So $d_{KL}(h||e^{-s(\cdot)}) \leq \kappa$ is a linear constraint

Snapshot of theory

True density $h^0 = e^{-s^0(\cdot)}$, with $s^0 \in \text{e-spl}^p(m)$

The following holds for estimator h^n:

Consistency
$h^n \to h^0$ as $n \to \infty$ uniformly (possibly except on m) w.p.1

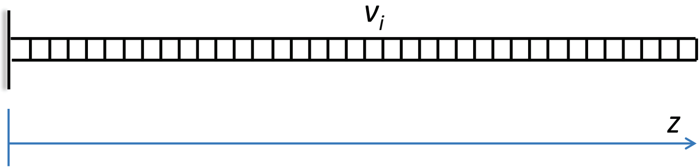
Asymptotics
For $x \in [m_0, m_N]$ and large n, $h^n(x)$ approx. normal w/mean $h^0(x)$

Rate of convergence
$n^{1/2} d_{KL}(h_\epsilon^{0,n} || h_\epsilon^n) = O_p(1)$

Example: Bar with random Young's modulus

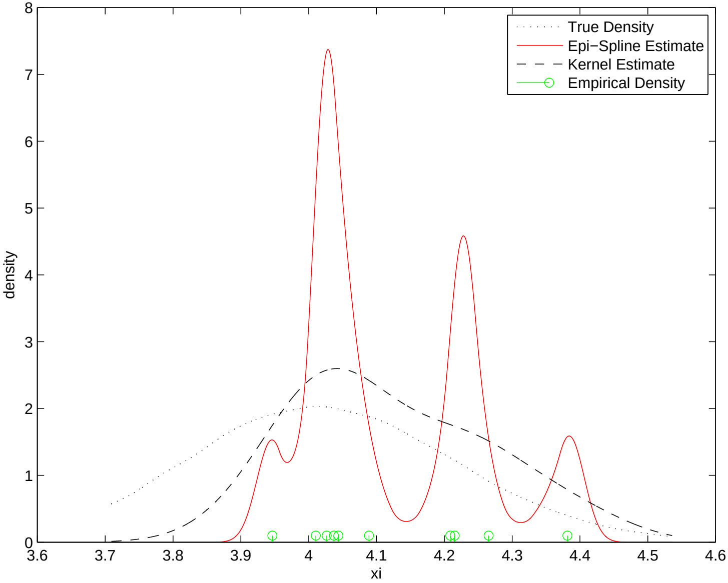
$$(E(z, \mathbf{v})u'(z))' = 0, \quad z \in (0, 1), \quad u(0) = u_0, E(1, \mathbf{v})u'(1) = 1$$

- $E(z, \mathbf{v})$ = Young's modulus at z under realization $\mathbf{v} = (v_1, v_2, ..., v_{100})$
- $E(z, \mathbf{v}) = v_i$ on $((i-1)/100, i/100)$ (piecewise constant)
- Components of $\mathbf{V}$ uniformly distributed on [0.1 0.5], independent



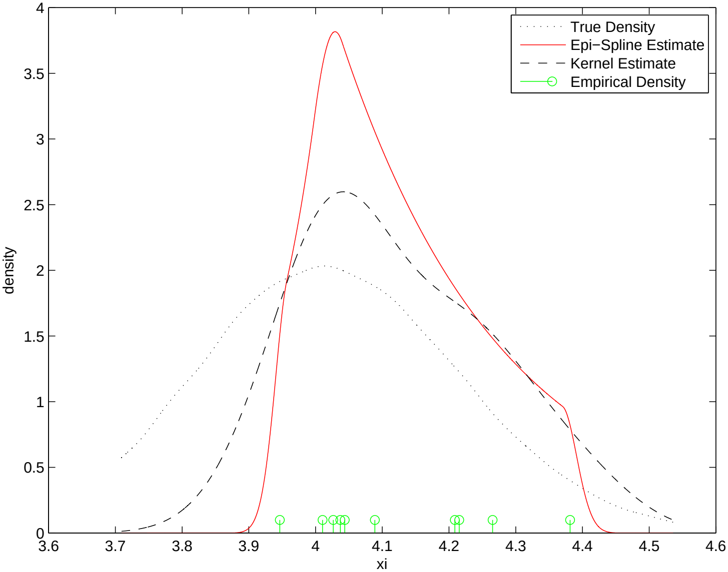
Bar: Density of endpoint displacement

Sample size 10; continuously differentiable



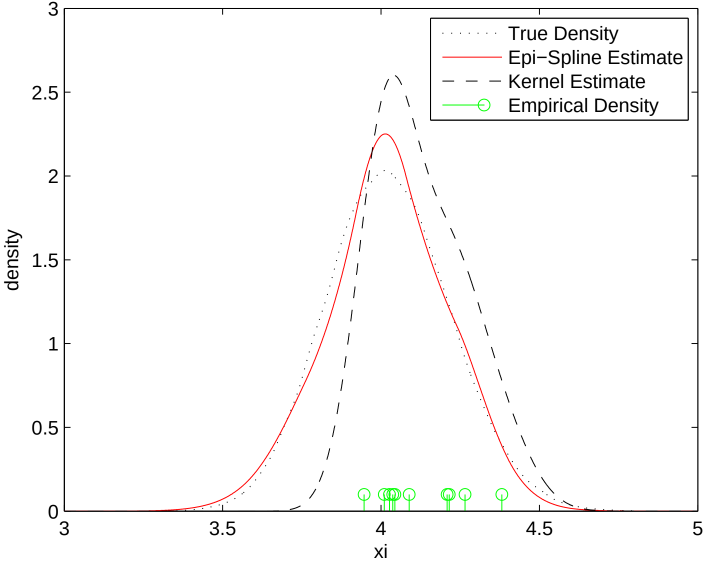
Bar: Density of endpoint displacement

Also unimodal



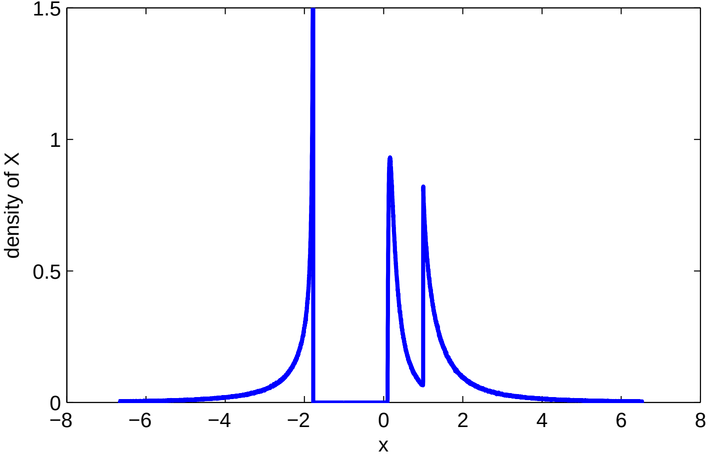
Bar: Density of endpoint displacement

Also support bounds and Kullback-Leibler constraint



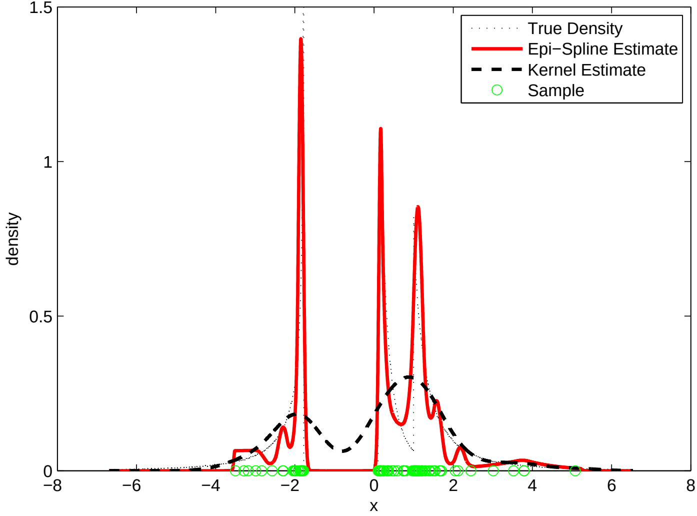
Back to 2-dof dynamical system

Recall the true density of the response



Dynamical system: Density of response

Sample size 100; continuously differentiable; "unimodal" tails



Dynamical System: Gradient information

Gradient information for bijective $G : \boldsymbol{R} \to \boldsymbol{R}$
Recall: If $X = G(V)$, then

$$h_X(x) = h_V(G^{-1}(x))/|G'(G^{-1}(x))|$$

Present context *without* a bijection and data $x^i = G(v^i), G'(v^i)$:

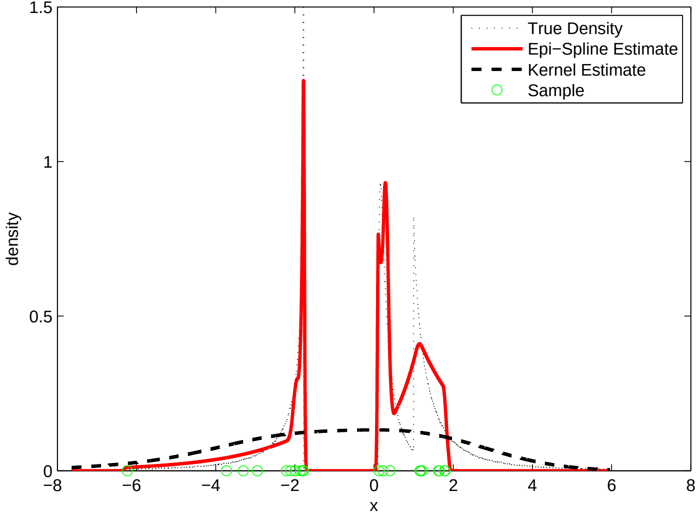
$$h^n(x^i) = e^{-\langle \mathbf{c}(x^i), \mathbf{r} \rangle} \geq \frac{h_V(v^i)}{|G'(v^i)|}$$

$$\langle \mathbf{c}(x^i), \mathbf{r} \rangle \leq -\log \frac{h_V(v^i)}{|G'(v^i)|}$$

Value of pdf bounded from below at x^i

Dynamical System: Gradient information

Sample size 20



Summary

- Exponential epi-splines offer a tractable class of density estimators
- Incorporate soft information by means of constraints
- Application to uncertainty quantification
- Extensions to response surface, regression curve, multivariate density estimation, and many other curve fitting problems

References

- R. & Wets, "Nonparametric density estimation via exponential epi-splines: fusion of soft and hard information," in review
- Singham, R., & Wets, "Density estimation of simulation output using exponential epi-splines," *Proceedings of the Winter Simulation Conference*, 2013
- R., Sukumar, & Wets "Uncertainty quantification using exponential epi-splines," *Proceedings of the International Conference on Structural Safety and Reliability*, 2013
- R. & Wets, "Epi-splines and exponential epi-splines: Pliable approximation tools," Naval Postgraduate School, Monterey, CA, 2013